

О МАТЕМАТИЧЕСКОМ МОДЕЛИРОВАНИИ ЛЕКСИКО- СЕМАНТИЧЕСКОЙ СИСТЕМЫ ЯЗЫКА

Иван Касабов

Одна из наиболее характерных черт современного состояния науки о языке в контексте достижения других наук — это усиленное стремление к большей точности и математизации исследования с целью прямой верификации и непосредственного применения теории в практике как критерия адекватности описываемого объекта. Эта большая точность и математизация обычно воспринимаются буквально, что приводит к попыткам свести описание языка к количественным отношениям посредством частотно-статистических или формально-логических исследований, имеющих при определенных целях свое основание.

Характер языка как сложный объект исследования нуждается в качественно новом понимании требований о точности и математизации с учетом задач, стоящих перед современной лингвистикой: дать полное описание по крайней мере одного естественного языка, который, по возможности, работал бы сам, т.е. через машину. Такое понимание исследования выражено путем принципа математического моделирования, который состоит в утверждении того, что объект можно считать доказанным тогда и лишь тогда, когда будет создан алгоритм его описания, его практическая модель как связь между идеальной теорией и реальным объектом и, соответственно, как доказательство адекватности теории. Этот принцип изложен и обоснован в области лингвистики, в частности лексикографии, Карауловым и Денисовым.

Известно, что, чтобы такая модель работала, она нуждается в определенном /конечном, сколь бы оно ни было большим/ количестве единиц и в нескольких однозначных правилах для их аранжировки. Ясно, что эти единицы должны

быть структурированы в довольно строгой системе. Чтобы претендовать на адекватность сложных реальных языковых фактов эта система и сама модель должны иметь достаточную представительную силу для данного языка в целостности. Сложность и практически бесконечное число единиц естественного языка требуют их искусственного ограничения в центральной части системы, согласно требования о необходимости и достаточности соблюдения специфики объекта. Это ограничение производится за счет полноты и подробности описания. За счет этих недостатков готовая же /и особенно действующая/ модель обладает большой эвристической силой: открывает новые закономерности и типологические характеристики системы.

Так как на современном этапе развития лингвистика все еще далеко от полного описания данного языка /тем более и языков/ и так как она далеко ушла в области грамматики, то труднейшей для моделирования частью языка является его семантика и прежде всего /или же, по крайней мере, эксплицитно/ его лексикальная семантика. Для моделирования этой области необходимо составление словаря — минимума как теоретического конструкта — репрезента лексики в целом.

Такой словарь — минимум может быть адекватный репрезент лексико-семантической системы болгарского языка в целостности так как он как словарь — конструкт представляет ее в системе и тем более в ее ядерной части и кроме того показывает отношения между единицами системы, т.е. ее структуру.

Результаты работы над таким словарем — минимумом: определение минимального /необходимого и достаточного/ корпуса из около 856 слов как заглавных единиц, который, кроме того, что охватывает и основную — ядерную часть лексико-семантической системы, может также выполнять функцию метаязыка для семантического описания всей системы, а также ее организация в лексико-семантических полях является доказательством семантической теории и базой для алгоритмизации лексикографической работы.